DOCUMENT RESUME

ED 128 390 TM 005 544

AUTHOR Polemeni, Anthony J.
TITLE Security in a Citywide Testing Program. NCME Measurement in Education, Vol. 6, No. 3, Summer 1975.
INSTITUTION National Council on Measurement in Education, East Lansing, Mich.
PUB DATE 75
NOTE 6p.
AVAILABLE FROM National Council on Measurement in Education, Office of Evaluation Services, Michigan State University, East Lansing, Michigan 48823 (Subscription rate; $5.00 per year; single copies $0.50 each in quantities of 25 or more, or $1.50 for a single issue)

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Achievement Tests; Cheating; *City Wide Programs; Confidentiality; Elementary Education; Reading Tests; Standardized Tests; *Testing Problems; *Testing Programs
IDENTIFIERS New York (New York); *New York City Reading Test; *Test Security

ABSTRACT

In April 1974, allegations were made that students, teachers, and the general public had access to the New York Citywide Reading Test prior to its administration, and the results, therefore, were invalid. In the face of these allegations, New York City developed a strategy for the administration of a secure test: a test never before available in the marketplace, and never before administered except for norming purposes. This document includes a step-by-step description of the procedures followed by the Office of Educational Evaluation in New York City. (BW)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED128390

Vol. 6, No. 3 NCME Summer 1975

measurement in education

A SERIES OF SPECIAL REPORTS OF THE NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICRO FICHE ONLY HAS BEEN GRANTED BY

NCME

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

Security in a Citywide Testing Program

Anthony J. Polemeni
New York City Board of Education

In April 1974, a furor erupted over the administration of the New York Citywide Reading Test and, as a result, the entire testing program had to be restructured. All students in grades two through nine in the public schools had taken the test according to a mandate of the New York State Legislature. Unfortunately, copies of the tests had fallen into the hands of newscasters and newspaper reporters prior to the administration of the test. The allegation was made that students, teachers, and parents also had prior access to the tests and the results, therefore, were invalid. As a consequence of all this, an investigation was launched into what were termed irregularities in the testing program. It was determined that in a few schools the actual test booklets had been used for coaching purposes and, while the overall impact had no perceptible influence on the citywide mean grade scores, public confidence in the use of commercially available standardized tests was effectively destroyed.

ABOUT THIS REPORT

Of the many issues currently involved in the attacks on standardized testing, test security is one of major importance, one which cannot be minimized. Particularly when viewed in relationship to the issues of reliability and validity does the use of a "secure test" take on added significance.

Dr. Anthony Polemeni finds himself in the sometimes unenviable position of being Director of the Office of Educational Evaluation for the city of New York, the nation's largest urban school system. In a city the size of New York, the potential problems border on the incredible. While the situation in New York may be idiosyncratic due to the enormous size and complexity of the system, certainly many of the points raised in this discussion are comparable to other situations where the issue of test security is of concern.

Dr. Polemeni has done extensive work in areas related to testing. He is well-published and has participated as a speaker at many professional meetings.

PSR

FROM THE EDITOR

Yes, it is May, 1976, and indeed, you are receiving the Summer, 1975 issue of *Measurement in Education*. I thank you for your patience and would like to take this opportunity to assure you that every attempt will be made to bring up to date your collection of ME, hopefully by the end of this calendar year.

At this time, let me also invite the readership to communicate directly with me pertaining to possible topics for consideration. Thanks again for your perseverence.

PSR

TM005544

". . . the notion of one's placement in the ranking being depressed through chicanery on the part of another would be most offensive."

Implications of the Problem

The situation was grave for three reasons: In the first place, the results of the Citywide Reading Test are used for the placement of pupils in compensatory and special education programs, and as one basis for the retention and promotion of pupils — a matter of tremendous concern to parents. Secondly, the Citywide Test results are used to rank all schools in the City of New York on the basis of reading performance. Obviously, there is a good deal of pride involved on the part of teachers and principals within each school, and the notion of one's placement in the ranking being depressed through chicanery on the part of another would be most offensive. Finally, but importantly, the reputation of 60,000 New York City public school teachers was being maligned because six or twelve of their number had acted foolishly.

In the face of these problems, New York City had only two options: Scuttle the Citywide Testing Program altogether, or develop a strategy for the administration of a secure test — a test never before available in the marketplace, and never before administered except for norming purposes.

"One assistant principal had been demoted, and several teachers had been officially reprimanded as a result of the scandal."

The dilemma gave rise to a series of high-level conferences to ensure that the matter be handled to the satisfaction of everyone involved. The serious nature of the problem was recognized: One assistant principal had been demoted, and several teachers had been officially reprimanded as a result of the scandal. No one wanted a repetition. In the final analysis, since a Citywide Reading Test score is necessary for a variety of purposes — including evaluation, allocation of funds, and administrative decision making — New York City chose to go with a secure testing program. It was understood, universally, that all procedures had to be so carefully defined that there could be no hint of improper practices. Such was the program that was developed in New York City.

A Step-By-Step Description

Since that time, several of the major cities in the United States have contacted New York City because they were encountering the same problems and wanted to know how New York had set up its program to ensure against irregularity, and allegation of irregularity. Since the replies were sketchy at best, and since increasing numbers of school systems throughout the country can anticipate similar problems, it was felt that a do-it-yourself-kit for security in a Citywide Testing Program might find a responsive readership. Such is the purpose of this article, and what follows is a step-by-step description of what was done by the Office of Educational Evaluation in New York City:

1. An application for pre-qualification as a bidder on the New York Citywide Testing Program was sent to 38 of the largest test publishing companies in the United States. Included in the documentation sent to the publishers were the general requirements for the tests, the answer documents, and the manuals. One stipulation of the pre-qualifying application read as follows: "The test shall be 'secure' in that it shall not be, nor ever have been, available to the public." To ensure against charges of favoritism, at the same time that the applications were sent to the 38 publishers, a public advertisement was placed in the City Record soliciting bids on the secure reading test.
2. In all, seven replies were received. Of these, five said they could not meet the requirements and specifications. One company said they had a secure test available, but the norms would not be available before September, 1975. This would have been too late to meet Office of Educational Evaluation time lines. Only one company replied that it had a secure standardized test available and normed, and could meet all stipulated requirements.
3. The Director of the Office of Educational Evaluation, the Coordinator of Citywide Testing, and a specialist in the New York City reading curriculum met with the publishers of the test to ascertain that the test was valid for New York City pupils, and that its reliability was acceptable. The tests were brought to the meeting by the publisher's representative, examined by Board of Education personnel, and removed by the publisher's representative.
4. At no time prior to the actual delivery of the tests by the publisher to the district depositories did any official or staff member of the Board of

Education keep a copy of the test in his/her possession. The purpose of this precaution was to ensure that should a leak occur it would be the responsibility of the publisher rather than of the Board of Education.

5. The title of the test was changed to the **New York City Reading Test** and it was reprinted by the publisher under maximum security procedures. These procedures included an actual count of each sheet of paper run through the printing press, and the shredding of all misprinted sheets.

"Compliance with all requests was maximal, since no one wanted a repetition of the furor that had accompanied the 1974 administration . . ."

---

6. Prior to the delivery of the tests to the districts, the Community School Superintendent within each district was required to select a depository to hold the testing materials for all schools within that district. It was made abundantly clear that security of the materials during the time they were in the district depository was the responsibility of the Community Superintendent and that depository, therefore, must be kept locked or guarded at all times. Compliance with all requests was maximal, since no one wanted a repetition of the furor that had accompanied the 1974 administration of the testing.
7. After a depository had been selected for each district, a staff member of the Office of Educational Evaluation visited each one to confirm that it was in fact, secure, and that it was large enough to accommodate the materials and the personnel to distribute them. It might be noted, too, that all depositories had to be on the ground floor, or accessible by freight elevator, in order that the trucker not be delayed in his schedule. (The entire delivery to the 32 districts, for the 1000 schools, had to be made in two days in order that there be minimum opportunity for the booklets to go astray.)
8. To oversee the depositories, each district provided two people (in most cases, the reading coordinator and the math coordinator) and the Office of Educational Evaluation provided one staff member. The function of these personnel, in the depositories, was to check the exact amount of materials delivered by the trucker, and to distribute these materials to school representatives.

9. As the school representatives arrived to pick up the materials, they had to present a "School Register As Of Date Of Testing." The purpose of this was to ensure that each school received no more materials than were necessary at the time. (The bulk printing order had been based on the October registers, and these figures had changed by April).
10. Once the material for a school had been picked up, and until that material was actually returned

to the test depository following completion of the test, its security was the responsibility of the school principal.

11. The test materials were picked up by the schools one day prior to the date set for test administration. This was necessary in order that there be time for distribution of the material to the teachers, and time for the teacher to fill in the identification grids. In most cases, the principal called a special staff conference on the afternoon of the day prior to test administration so that teachers might be properly instructed in the use, coding, packaging, and labeling of the materials.

12. All tests, in all second through ninth grade classes, in all public schools in New York City, were administered on the same day. No exceptions were permitted. Those students who were absent on the day of the test were retested at a later date with a different form of the test. The scores of these retested students, while they were given to teachers for classroom use, were not entered in the statistical analysis of the Citywide Reading Survey.

13. During the time of test administration, staff members of the Office of Educational Evaluation made unannounced visits to approximately 75 schools throughout the city. These visits were unannounced only in the sense that no school knew whether or not it would be visited; all schools had been put on notice that such monitoring would occur on a random basis. No representative of the Office of Educational Evaluation recorded any untoward incident during these visits to the schools.

"No representative of the Office of Educational Evaluation recorded any untoward incident..."

---

14. Every teacher had to submit an answer document for every student on register as of the testing date. The answer document had to be coded as "tested," "absent," or "excused as non-English speaking." A student could be excused as non-English speaking "... who in the opinion of the school cannot reasonably be expected to read or understand test content because of language-related difficulties." Pupils in retarded mental development, junior guidance, health conservation, or visually handicapped classes were not included in the testing program at all since they were not on regular class register.

REPORTS AVAILABLE

Back issues of *Measurement in Education* are available at 50¢ each in quantities of 25 or more for a single issue.

Vol. 1, No. 1	*Helping Teachers Use Tests* by Robert L. Thorndike
No. 2	*Interpreting Achievement Profiles — Uses and Warnings* by Eric F. Gardner
No. 3	*Mastery Learning and Mastery Testing* by Samuel T. Mayo
No. 4	*On Reporting Test Results to Community Groups* by Alden W. Badal & Edwin F. Larsen
Vol. 2, No. 1	*National Assessment Says* by Frank B. Womer
No. 2	*The PLAN System for Individualizing Education* by John C. Flanagan
No. 3	*Measurement Aspects of Performance Contracting* by Richard E. Schutz
No. 4	*The History of Grading Practices* by Louise Witmer Cureton
Vol. 3, No. 1	*Using Your Achievement Test Score Reports* by Edwin Gary Joselyn & Jack C. Merwin
No. 2	*An Item Analysis Service for Teachers* by Willard G. Warrington
No. 3	*On the Reliability of Ratings of Essay Examinations* by William E. Coffman
No. 4	*Criterion-Referenced Testing in the Classroom* by Peter W. Airasian and George F. Madaus
Vo. 4, No. 1	*Goals and Objectives in Planning and Evaluation: A Second Generation* by Victor W. Doherty and Walter E. Hathaway
No. 2	*Career Maturity* by John O. Crites
No. 3	*Assessing Educational Achievement in the Affective Domain* by Ralph W. Tyler
No. 4	*The National Test-Equating Study in Reading (The Anchor Test Study)* by Richard M. Jaeger
Vol. 5, No. 1	*The Tangled Web* by Fred F. Harcleroad
No. 2	*A Moratorium? What Kind?* by William E. Coffman
No. 3	*Evaluators, Educators, and the Publics: A Detente?* by William A. Mehrens
No. 4	*Shall We Get Rid of Grades?* by Robert L. Ebel
Vol. 6, No. 1	*On Evaluating a Project: Some Practical Suggestions* by John W. Wick
No. 2	*Measuring Reading Achievement: A Case for Criterion — Referenced Testing and Accountability* by S. Jay Samuels and Glenace E. Edwall

15. Immediately following the test administration each teacher wrapped and labeled (separately) the answer documents, the used test booklets, and the unused test booklets and teacher manuals. These packages were then sent to the principal's office. No remnant of the testing materials was to remain in the classroom of any teacher.

16. When the testing materials from each classroom had been gathered in the principal's office, they were returned to the district test depository where a receipt was issued. Again, no remnant of the testing program was to remain in any school.

17. On the first or second day following the test date (and during which time the depositories remained guarded or locked), the materials were picked up by the trucker — in the presence of an Office of Educational Evaluation representative — and shipped to the scoring centers.

18. While the tests were being scored, the test publisher began work on the development of a parallel form of the test for administration in 1976. That test has been administered under exactly the same security procedures described above since, as a result of the security procedures, there was not a single allegation of irregularity during or following the entire testing program.

Conclusion

In summary then, New York City, faced with the problem of developing a secure citywide test, developed a strategy and solved the problem. This was fine for New York City. But now a very pointed question: Supposing other large cities — or, indeed, entire states — want to replicate the New York City strategy. Where do all the "secure" tests come from? This is a question for the major test publishers to answer. It is likely, in the light of New York City experience, that they have already begun working on the answer. Test publishers are in business to make money; they must provide what the consumer demands. If the questions addressed to New York City (which this article has attempted to answer) are a portent, then ever-increasing numbers of consumers will be demanding secure tests.

The consumer, for his part, must be willing to pay a price for the security of his testing program. New York City, for example, paid $129,300.00 in developmental costs for the 1976 version of its Citywide Reading Test. This is a lot of money at any time; it is a tremendous amount of money in this day of shrinking educational budgets.

Perhaps what is needed is the formation of an ad-hoc "think-tank" composed of Chief School Officers, Heads of Evaluation, and fiscal and technical experts from the major test publishing companies throughout the United States. If citywide testing is to continue, then educators, parents, and students have enough to worry about in terms of validity and culture-fairness. They should not have the additional concern that test results are invalid because the testing program itself was not secure.